

# **INTERCONNECTION ARCHITECTURE FOR MANAGING MULTIPLE LOW BANDWIDTH CONNECTIONS OVER A HIGH BANDWIDTH LINK**

**Publication number:** JP2002540701 (T)

**Publication date:** 2002-11-26

**Inventor(s):**

**Applicant(s):**

**Classification:**

- **international:** **G06F13/40; H04L12/46; G06F13/40; H04L12/46; (IPC1-7): H04L12/46**

- **European:** **G06F13/40D2**

**Application number:** JP20000608267T 20000329

**Priority number(s):** US19990126825P 19990330; US20000502947 20000211; WO2000US08307 20000329

**Also published as:**

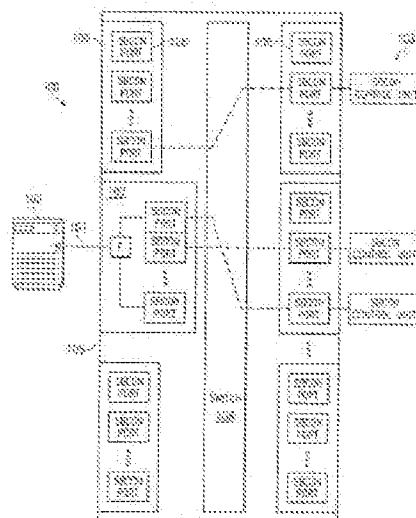
WO0058842 (A1)  
US2004174867 (A1)  
US6975623 (B2)  
US6728803 (B1)  
EP1183608 (A1)

[more >>](#)

Abstract not available for JP 2002540701 (T)

Abstract of corresponding document: **WO 0058842 (A1)**

A data communication architecture (100) including a plurality of devices (103) having input/output (I/O) ports supporting communication at a first rate and a data processor (302) having a number of I/O ports where each I/O port support data communication at a second data rate. The second data rate is at least double the first data rate. A communication link (101) coupled to one of the data processor I/O ports supports the second data rate. A bridge device (107) is coupled to the communication link and to the I/O ports of the plurality of devices. The bridge device translates the communication link at the second data rate to a plurality of communication links at the first data rate, where the plurality of communication links at the first data rate are substantially independent of each other.



Data supplied from the **espacenet** database — Worldwide

(19) 日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11) 特許出願公表番号  
特表2002-540701  
(P2002-540701A)

(43) 公表日 平成14年11月26日 (2002. 11. 26)

(51) Int.Cl.<sup>7</sup>  
H 0 4 L 12/46

識別記号

F I  
H 0 4 L 12/46

デマコト\* (参考)

Z 5 K 0 3 3

審査請求 有 予備審査請求 有 (全 44 頁)

(21) 出願番号 特願2000-608267(P2000-608267)  
(86) (22) 出願日 平成12年3月29日 (2000. 3. 29)  
(85) 翻訳文提出日 平成13年10月1日 (2001. 10. 1)  
(86) 国際出願番号 PCT/US 00/08307  
(87) 国際公開番号 WO 00/58842  
(87) 国際公開日 平成12年10月5日 (2000. 10. 5)  
(31) 優先権主張番号 60/126, 825  
(32) 優先日 平成11年3月30日 (1999. 3. 30)  
(33) 優先権主張国 米国 (US)  
(31) 優先権主張番号 09/502, 947  
(32) 優先日 平成12年2月11日 (2000. 2. 11)  
(33) 優先権主張国 米国 (US)

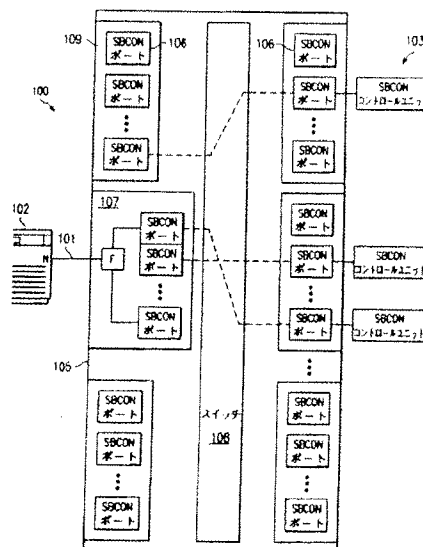
(71) 出願人 マックデータ・コーポレーション  
McDATA CORPORATION  
アメリカ合衆国コロラド州80021, ブルームフィールド, インターロッケン・パークウェイ 310  
310 Interlocken Parkway, Broomfield, Colorado 80021, United States of America  
(74) 代理人 弁理士 浅村 皓 (外3名)

最終頁に続く

(54) 【発明の名称】 高帯域幅リンクを介した多数の低帯域幅接続を管理する相互接続アーキテクチャ

(57) 【要約】

第1のレートの通信をサポートする入出力 (I/O) ポートを有する複数の装置 (103) といくつかの I/O ポートを有し各 I/O ポートが第2のデータレートのデータ通信をサポートするデータプロセッサ (302) とを含むデータ通信アーキテクチャ (100)。第2のデータレートは第1のデータレートの少なくとも2倍である。データプロセッサ I/O ポートの1つに接続された通信リンク (101) が第2のデータレートをサポートする。ブリッジ装置 (107) が通信リンクおよび複数の装置の I/O ポートに接続されている。ブリッジ装置は第2のデータレートの通信リンクを第1のデータレートの複数の通信リンクへ変換し、第1のデータレートの複数の通信リンクは互いに実質的に独立している。



**【特許請求の範囲】**

【請求項1】 データ通信アーキテクチャであって、

第1のデータレートの通信をサポートする入出力（I/O）ポートを有する複数の装置と、

いくつかのI/Oポートを有するデータプロセッサであって、各I/Oポートが第2のデータレートの通信をサポートし、第2のデータレートは第1のデータレートの少なくとも2倍であるデータプロセッサと、

データプロセッサI/Oポートの1つに接続されて第2のデータレートをサポートする通信リンクと、

通信リンクおよび複数の装置のI/Oポートに接続されたブリッジ装置であって、第2のデータレートの通信リンクを第1のデータレートの複数の通信リンクへ変換するブリッジ装置と、

を含むデータ通信アーキテクチャ。

【請求項2】 請求項1記載のデータ通信アーキテクチャであって、第1のデータレートの複数の通信リンクは互いに実質的に独立しているデータ通信アーキテクチャ。

【請求項3】 請求項1記載のデータ通信アーキテクチャであって、複数の装置のI/Oポートは回線交換通信をサポートし、

データプロセッサのI/Oポートはパケット交換通信をサポートする、データ通信アーキテクチャ。

【請求項4】 請求項1記載のデータ通信アーキテクチャであって、複数の装置のI/OポートはSBCONをサポートするデータ通信アーキテクチャ。

【請求項5】 請求項1記載のデータ通信アーキテクチャであって、データプロセッサのI/Oポートはファイバチャネルをサポートする、データ通信アーキテクチャ。

【請求項6】 請求項1記載のデータ通信アーキテクチャであって、データプロセッサは複数のメッセージを発生し、各メッセージが複数の装置の特定の1つを有するバーチャルチャネルに対応するデータ通信アーキテクチャ。

【請求項7】 請求項6記載のデータ通信アーキテクチャであって、さら

に、複数のバーチャルチャネルを多重化して単一通信リンクとするように接続されたデータプロセッサ内のマルチプレクサを含むデータ通信アーキテクチャ。

【請求項8】 請求項1記載のデータ通信アーキテクチャであって、ブリッジは、さらに、

フレームプロセッサとして動作するフロントエンドと、

通信リンクを初期化するように動作するローカルデータプロセッサと、

複数の装置の各々に対するリンクコントローラを含むリンクコントローラユニットであって、各リンクコントローラが複数の装置の対応する1つとの通信チャネルをサポートするリンクコントローラユニットと、

を含むデータ通信アーキテクチャ。

【請求項9】 請求項8記載のデータ通信アーキテクチャであって、さらに、

データプロセッサへ交換クレジットメッセージを発生するように動作するローカルプロセッサ内の第1の交換クレジット機構と、

交換クレジットメッセージを受信するように動作するデータプロセッサ内の第2の交換クレジット機構と、

を含むデータ通信アーキテクチャ。

【請求項10】 請求項8記載のデータ通信アーキテクチャであって、第1の交換クレジット機構はリンクコントローラユニット内のいくつかの動作リンクコントローラに応答するデータ通信アーキテクチャ。

【請求項11】 請求項8記載のデータ通信アーキテクチャであって、第1の交換クレジット機構は動作リンクコントローラに関連する複数の装置の組合せデータレートに応答するデータ通信アーキテクチャ。

【請求項12】 データ接続の管理方法であって、

第1のメッセージを発生するステップと、

第1のメッセージを符号化して第1のデータグラム (datagram) とするステップであって、第1のデータグラムはパケット交換リンクに使用されるメタデータ (meta-data) を含むステップと、

パケット交換リンクを介してデータグラムを運ぶステップと、

リンクに接続された中間データトランスポート機構内のパケット交換リンクからデータグラムを受信するステップと、

中間データトランスポート機構内にメタデータを格納するステップと、

第1のメッセージを再符号化して第2のデータグラムとするステップであって、第2のデータグラムは回線交換リンクに使用されるメタデータを含むステップと、

回線交換リンクを介して第2のデータグラムを運ぶステップと、  
を含む方法。

【請求項13】 請求項12記載の方法であって、さらに、  
回線交換リンクから第2のデータグラムを受信し、  
第2のデータグラムに応答して第2のメッセージを発生し、  
第2のメッセージを符号化して回線交換リンクに使用されるメタデータを含む第3のデータグラムとし、

回線交換リンクを介して中間データトランスポート機構へ第3のデータグラムを運び、

第2のメッセージを再符号化して第4のデータグラムとし、第4のデータグラムはパケット交換リンクに使用される格納されたメタデータからコピーされたメタデータを含む、

方法。

【請求項14】 請求項12記載の方法であって、第1のメッセージはメッセージが属する論理交換を一意的に識別する値を保持する交換識別フィールドを含む方法。

【請求項15】 請求項14記載の方法であって、さらに、  
中間データトランスポート機構内のパケット交換リンクから第1のデータグラムを受信した後で、論理交換の持続時間中交換識別値を特定の回線交換リンクに結合する、ステップを含む方法。

【請求項16】 通信リンクに対するブリッジ回路であって、  
全二重パケット交換リンクをサポートするパケット交換側と、  
いくつかの全二重回線交換リンクをサポートする回線交換側と、

パケット交換フレームを特定の1つの回線交換リンクに結合する論理結合記述を格納する記憶スペースを有するブリッジ回路内の結合機構と、  
を含むブリッジ回路。

【請求項17】 請求項16記載のブリッジ回路であって、ブリッジ回路はパケット交換リンクを介して受信したパケット交換フレーム内に表示された論理交換を識別し、論理交換の持続時間中ずっと論理結合を維持するブリッジ回路。

【請求項18】 請求項16記載のブリッジ回路であって、結合機構は、さらに、

受信したパケット交換フレームから選択されたヘッダー情報を保持する記憶構造と、

格納されたヘッダー情報を使用して受信した回線交換フレームをパケット交換フレームへ再フォーマット化するフレーム発生器と、

を含むブリッジ回路。

【請求項19】 通信リンクの操作方法であって、

高帯域幅接続および複数の低帯域幅接続をサポートするブリッジユニットを提供するステップと、

低帯域幅接続の動作性（operability）を検証するステップと、

動作可能な低帯域幅接続数に基づいて交換機クレジット値を決定するステップと、

高帯域幅接続を介してクレジット値を含むメッセージを発行するステップと、  
高帯域幅接続に接続された任意の装置に、その装置から高帯域幅接続を介してブリッジユニットにより通信が受け入れられる前に、少なくとも1つの交換機クレジットを持つことを要求するステップと、

を含む方法。

【請求項20】 請求項19記載の方法であって、検証はブリッジ回路の初期化中に実施される方法。

【請求項21】 請求項19記載の方法であって、検証は実行時間に動的に実施される方法。

【請求項22】 請求項19記載の方法であって、さらに、

高帯域幅接続に接続された装置からブリッジユニット内のメッセージを受信するステップであって、メッセージは交換クレジットおよび交換識別子を有するステップと、

交換識別子を低帯域幅接続の選択された1つに結合することにより論理交換をオープンとするステップと、

同じ交換識別子を有するブリッジユニットにより受信された後続メッセージをその交換識別子に結合される低帯域幅接続へルーティングするステップと、  
を含む方法。

## 【発明の詳細な説明】

## 【0001】

## (発明の背景)

## 1. (発明の分野)

本発明は、一般的にデータ通信に関し、特に帯域幅のより高い単一の通信チャネルを介した多数の低帯域幅接続を管理するシステム、方法およびアーキテクチャに関する。

## 【0002】

## (関連する背景)

企業の計算網は高速通信チャネルによりリンクされる地理的に分散された計算リソースの集合により形成される。典型的には、1つ以上のメインフレームコンピュータがバルクデータ処理を行うのに使用され、他のノードは特殊化された機能のため使用される。1つの例は通信チャネルや網によりメインフレームプロセッサに接続される“記憶ファーム”内に大容量記憶装置が実現される記憶エリア網(SAN)である。

## 【0003】

ここで使用される、通信“チャネル”は通信装置間に直接すなわち交換ポイントツーポイント接続を提供する。回線交換チャネルは典型的にハードウェア集約的であり、チャネル管理に必要な僅かなオーバーヘッドでデータを高速で運ぶ。通常回線交換接続はデータが転送されない場合でも確立されたままであり、したがって帯域幅が浪費され、しかも時分割多重化等の多重化技術を介して多数のユーザをサポートすることができる。

## 【0004】

一方、パケット交換網では網媒体および可変長パケットを使用する利用可能な帯域幅をユーザがダイナミックに共用することができる。パケット交換網は回線交換通信に比べてより効率的で柔軟なデータ転送により特徴づけられる。パケット交換通信は行き先に達するまでさまざまな網構成要素間でパケットを交換できるようにするアドレッシング情報を各パケットに加えることによりオーバーヘッドを増加する。

## 【0005】

高帯域幅長距離通信を実現するための初期の試みには回線交換技術が使用された。広く使用されたこのような技術の例はSBCON (single byte command code sets connection) アーキテクチャである。SBCONは米国国内標準協会 (ANSI) 標準X3.296-1997 “Information Technology—Single-Byte Command Code Sets Connection (SBCON) Architecture” により標準化されている。ANSIドキュメントX3.296-1997には光ファイバリンク、交換ポイントツーポイントトポロジー、および高帯域幅、高性能長距離情報交換について明記する入出力 (I/O) および相互接続アーキテクチャが記述されている。ここで使用されるSBCONはIBMから提供される企業システム接続 (ESCON) アーキテクチャ、その他のSBCONのバリエーションだけでなく標準SBCONアーキテクチャに関係している。本発明の目的に対して、これらのバリエーションはSBCONと同等と見なされる。

## 【0006】

SBCONは最大200 Mbit/秒の全二重チャネルをサポートする。SBCONは分散アーキテクチャ内のメインフレームと記憶装置その他の周辺構成要素間の通信をサポートするために広く展開されている。したがって、SBCONアプリケーションおよび装置の設置されたかなりのベースが存在する。しかしながら、通信およびデータ処理および記憶技術の急速な発展により多くのSBCON施設は最適ではないものとされてしまっている。

## 【0007】

一般的に分散計算環境、特に、SANアプリケーションは装置間に次第に高速となる通信リンクを必要とする。従来のメインフレームアーキテクチャはメインフレームと他の装置間を接続するオペレーティングシステムで定義され、システムで限定された数 (例えば、256) の接続ポートをサポートする。データ処理速度の性能改善により次第にデータ集約的かつ速度集約的アプリケーションが生じてきている。データ転送に対する要求が増すにつれ、従来の通信技術のチャンネル当たり200 Mbit/秒の制限が制約的となっている。メインフレームオペレーティングシステムはより多くのポートを提供するように容易に変更できないため、増加したデータ転送に対する唯一の解決法は各チャンネルの帯域幅を増すことである

。

## 【0008】

ワークステーション、メインフレーム、スーパーコンピュータ、記憶装置その他周辺装置間の高帯域幅データ転送用の拡張可能で、柔軟性のある通信アーキテクチャとしてファイバーチャネルが開発されてきている。ファイバーチャネルは256 Mbits/秒（双方向）から2 Gbits/秒（双方向）までの多様な速度範囲で動作し、4 Gbit/秒の速度まで考えられる。銅および光通信媒体の両方に対して標準が規定されている。ファイバーチャネルはパケット交換および回線交換通信の両方の望ましい特徴を合わせ持っている。ファイバーチャネルは装置を接続する“ファブリック（fabric）”と呼ばれるアクティブで、インテリジェントな相互接続アーキテクチャを使用する。ファイバチャネルの物理的なインプリメンテーションはパケット交換であるが、ファブリックはノード間の専用バーチャル接続を含むサービスのさまざまなクラスをサポートして、さまざまなタイプのトラフィックの効率的な伝送を保証する。

## 【0009】

ファブリックは装置がファブリックへアクセスできるようにする、F\_Portsと呼ばれる、いくつかのポートを提供する。装置はその中に実現されたもしくはそれに関連するノードポート（N\_Port）を使用してF\_Portsに接続する。ファイバチャネルファブリックに接続するために、装置はファブリック接続を管理するノードポートすなわち（N\_Port）を含んでいる。N\_PortはファブリックポートすなわちF\_portを有するファブリックエレメント（例えば、スイッチ）への接続を確立する。ファブリックに取り付けられた装置はN\_PortとF\_port間の接続を管理するのに十分なだけインテリジェンスしか必要としない。ファブリックエレメントはルーティング、エラー検出および回復、および類似の管理機能を処理するためのインテリジェンスを含んでいる。

## 【0010】

スイッチは各F\_portがそれ自体とその付属システム間の単純なポイントツーポイント接続（point-to-point connection）を管理する多数のF\_portを有する装置である。各F\_portはサーバ、周辺装置、I/Oサブシステム、もしくはブリッ

ジに取り付けることができる。スイッチは1つのポートから接続要求を受信し、要求内に含まれるアドレス情報に基づいてもう1つのポートへの接続を自動的に確立する。多数の呼もしくはデータ転送がマルチポートファイバチャネルスイッチを介して同時に生じる。交換技術の重要な利点は、スイッチを介して一度接続が確立されると、その接続により提供される帯域幅は共用されない点において“非ブロッキング (non-blocking)” であることである。したがって、銅配線や光ファイバケーブルリング等の物理的接続リソースは、必要に応じて多数のユーザが物理的接続リソースにアクセスできるようにすることにより、より効率的に管理することができる。

#### 【0011】

ファイバチャネルは遥かに高帯域幅の接続技術を提供するが、SBCON装置のような過去から受け継いだ回線交換システムの大規模に設置されたベースはファイバチャネルファブリックに直接接続することができない。網内のメインフレームコンピュータへのファイバチャネルポートを提供することは実現可能ではあるが、ファブリックとのインターフェイスを更新もしくは置換する必要があるノード装置は数百台にもなることがある。その結果、ファイバチャネルにより提供されるより高速技術への移行が遅くなったり、場合によっては実現するのに非常に費用がかかるようになってきている。

#### 【0012】

SBCONおよびESCONトラフィックをファイバチャネルパケット内にカプセル化するすなわち埋め込む努力がなされてきている。これらの解決法はSBCONトラフィックをファイバチャネル通信媒体とコンパチブルとする。しかしながら、高帯域幅ファイバがSBCON装置にデータを供給している限り、ファイバチャネル通信リンクはSBCON装置が受け入れるのに等しい有効データレートでしか動作することができない。したがって、ファイバチャネルの多くの利点は過去から受け継いだSBCONもしくはESCON装置にアクセスする時に浪費される。したがって、低帯域幅接続を効率的に運ぶファイバチャネルのような高速通信リンクを可能にする接続アーキテクチャが必要とされている。

#### 【0013】

(発明の開示)

簡単に言えば、本発明は全二重パケット交換リンクをサポートするパケット交換側およびいくつかの全二重回線交換リンクをサポートする回線交換側を有する通信リンクに対するブリッジ回路を含んでいる。ブリッジ回路内の結合機構が論理結合記述を格納するデータ構造を維持する。論理結合記述はパケット交換フレームを回線交換リンクの特定の1つに結合する。

**【0014】**

また、本発明は第1のレートの通信をサポートする入出力(I/O)ポートを有する複数の装置および各々が第2のデータレートの通信をサポートするいくつかのI/Oポートを有するデータプロセッサを含むデータ通信アーキテクチャを含んでいる。第2のデータレートは第1のデータレートの少なくとも2倍である。データプロセッサI/Oポートの1つに接続された通信リンクは第2のデータレートをサポートする。ブリッジ装置が通信リンクおよび複数の装置のI/Oポートに接続されている。ブリッジ装置は第2のデータレートの通信リンクを複数の第1のデータレートの通信リンクへ変換し、複数の第1のデータレートの通信リンクは互いに実質的に独立している(すなわち、第1のデータレートのリンクは制御、シグナリング、もしくはデータ情報を共用する必要がない)。

**【0015】**

もう1つの面において、本発明は高帯域幅接続および複数の低帯域幅接続をサポートするブリッジユニットとの通信リンクを操作する方法を含んでいる。低帯域幅接続の操作性が検証され、操作可能な低帯域幅接続数に基づいて交換クレジット値(exchange credit value)が決定される。クレジット値を含むメッセージが高帯域幅接続で発行される。高帯域幅接続に接続されたいかなる装置も、その装置からの通信が高帯域幅接続を介してブリッジユニットより受け入れられる前に、少なくとも交換クレジットを有する必要がある。

**【0016】**(好ましい実施例の詳細な説明)

本発明は高帯域幅データ通信リンクを使用する低帯域幅装置との効率的な通信に有用なデータ通信アーキテクチャに向けられている。本発明はメインフレー

ムコンピュータと記憶装置およびプリンタ等の複数のシングルビットコマンドチャンネルコマンドセット接続（SBCON）共用周辺装置間でデータを運ぶのにファイバチャンネルが使用される特定の実施例に関して記述される。しかしながら、この特定例は、いかなる理由であれ、低帯域幅装置を高帯域幅通信リンクに接続したより一般的なデータ通信応用に容易に拡張することができる。本発明は低帯域幅装置が回線交換型通信チャンネルをサポートし、高帯域幅リンクがパケット交換技術をサポートする環境において特に有用であるが、他の応用も容易にお判りであろう。

#### 【0017】

図1は本発明が有用に利用される単純化された分散形計算環境100を示す。この環境100はメインフレームコンピュータ102をディレクタ105を介して、とりわけ、選択されたSBCON装置103とインタラクト（interact：対話）できるようにする。1つのメインフレームすなわち“ホスト”102しか図示されていないが、典型的な環境は各々が1つ以上のディレクタ105に接続されたFCリンクおよび／もしくはSBCONリンクを有することができる多数のホストを含むことができる。

#### 【0018】

計算環境100は、例えば、光ファイバケーブルにより1つ（以上の）ディレクタ105に接続されるSBCONコントロールユニット103等の複数の装置を含むこともできる。ディレクタ105は、図1の特定の例ではSBCONポートと表示されている、そのポート対106を動的に接続することができる。スイッチ108を介したこれらの接続の動的性質（dynamic nature）は図1に破線接続により示唆される。個別の各ポート106が任意所与の時間に1接続を行うことができる。

#### 【0019】

特定実施例のディレクタ105にはポートカード109をプラグインすることができるいくつかの拡張スロットが設けられている。各ポートカードは一群のポート106（例えば、8つ）を実現する回路および装置を含んでいる。このようにして、必要に応じて付加ポート106をプラグインすることによりディレクタ

105を拡張することができる。

#### 【0020】

図1の特定の実施例では、例えば、1つの拡張スロット内へプラグインすることにより、1つ以上のブリッジ装置107がディレクタ105に接続される。ブリッジ装置107はディレクタ105の拡張スロットとのコンパチブルなインターフェイスにより構成される。理想的には、ディレクタ105はポートコントローラ107により提供されるSBCONポート106を他の従来のSBCONポート106から識別することができない。しかしながら、従来のポートカードとは異なり、ブリッジ装置107はF\_portインターフェイスをファイバチャネルリンク101に提供する。特定の例では、1つのF\_portが8つのSBCONポートをサポートする。

#### 【0021】

ファイバチャネル（FC）リンク101は銅もしくは光ファイバ物理的接続技術もしくは任意の利用可能な同等技術等のファイバチャネルコンプライアントハードウェア（fiber channel compliant hardware）およびソフトウェアを使用して実現される。リンク101は図1に“N”で示すノードポート（N\_Port）と図1に“F”で示すファブリックポート（F\_Port）間のポイントツーポイントもしくはバーチャルポイントツーポイントリンクを含んでいる。単一のFCリンク101が例示されているが、典型的なシステムではメインフレーム102内の任意数の利用可能なポートを本発明を使用してブリッジ装置107に接続することができる。

#### 【0022】

メインフレーム102は高速データ処理マシンを含み、シングルプロセッサもしくはマルチプロセッサ装置として実現することができる。メインフレーム102はプロセッサがデータ処理およびソフトウェア命令実行のためにアクセスすることができるメモリ装置を含んでいる。メインフレーム102はローカル大容量記憶装置、入出力（I/O）装置、その他の利用可能な装置および特定アプリケーションのデータ実行ニーズを助ける周辺装置を含んでいる。本発明に従ったコンピュータプログラムプロダクト装置の一部はメモリおよびメインフレーム10

2に関連する大容量記憶装置内に格納され、メインフレーム102内のプロセッサにより実行される。メインフレーム102は任意の市販のもしくは特殊目的コンピュータコンポーネントおよび技術を使用して実現することができる。

#### 【0023】

SBCONコントロールユニット103は各々がプリンタ、大容量記憶装置、磁気ディスク装置、光ディスク記憶装置、テープ記憶装置、等の1つ以上の共用周辺装置を含んでいる。コントロールユニット103は分散形計算環境100内の任意タイプの周辺装置と置換できることをお判り願いたい。

#### 【0024】

ブリッジユニット107は最小200 Mbits/秒で動作し利用可能なコンポーネントに応じて1 Gbits/秒以上まで動作する全二重通信リンクをサポートする。各SBCONチャネルポート106が特定の例において200 Mbits/秒二重チャネルをサポートする。本質的に、ポートコントロール107は8つの200 Mbits/秒チャネルを結合して1 Gbits/秒リンクで送信するためのマルチプレクサ/デマルチプレクサ (MUX/DEMUX) ユニットとして機能する。

#### 【0025】

本発明に従って、ブリッジユニット107はSBCONインターフェイスの回線交換要求条件およびメインフレーム102とのパケット交換接続をサポートする。FC-PH仕様に規定されているようなFC-0からFC-1伝送およびシグナリングプロトコルを実現するのに任意の利用可能な機構および回路を使用することができる。本発明はパケット間の交換結合をポートコントロール107の回線交換側のメインフレーム102および個別の回路から実現する。本発明に従った交換結合によりメインフレーム102および各SBCONコントロールユニット103は特定の各通信交換に関する状態情報を維持することができる。この状態情報によりメインフレーム102から発生されたメッセージをSBCONコントロールユニット103により発生された応答と対とすることができ、その逆もできる。このようにして、比較的高速のファイバチャネル通信リンクを多重化して多数の低帯域幅（例えば、SBCON）チャネルをサポートするのに効率的に使用することができる。

#### 【0026】

本発明はメインフレーム102とポートコントロール107間の情報交換 (information exchange) の基本単位である“情報単位”の概念を使用する。情報単位はデータフレームやデータパケットの多くの特性を有するが、ファイバチャネルFC-4等のもう1つのトランスポートパケットのペイロード部に埋め込まれる。図2は本発明に従った典型的な情報単位201を示す。各情報単位201は、例えば、FC-2フレームとして実現される1つ以上のフレーム202を含んでいる。一般的に、FC-2フレーム202はフレーム開始(SOF) キャラクタ (start of frame character) およびフレーム終了(EOF) キャラクタ (end of frame character) により線引される (delineated)。各FC-2フレームはマルチビットFC-2フレームヘッダー、オプションFC-4ヘッダーおよび巡回冗長検査(CRC) フィールドだけでなくペイロードすなわちデータフィールド (図示せず) を含んでいる。

#### 【0027】

メインフレーム102はブリッジ装置107と協調して、各々が1つのSBCONリンクに対応する複数の交換を管理する。対応するSBCONリンクは受信したFCフレーム202のFC-4ヘッダー内に指定される。各交換はその持続時間中、指定されたSBCONリンクに結合される (bound)。SBCONリンクは図1に示すメインフレーム102上で実行するソフトウェアにより決定すなわち指定される。メインフレーム102は、そのSBCONリンクに向けられた後の全ての通信に交換IDが付けられるように (tagged with the exchange ID)、SBCONリンクと“交換ID”間の結合 (binding) を作り出す。本質的に、交換IDは各オープン交換を一意的に識別する2進値である。例えば、合計8つの交換が一時にオープンとなることがある場合には、各交換を一意的に識別するのに2進値“000”から“111”を使用することができる。各コマンドが特許特定のSBCONリンクを指定する限り、交換IDを選択するための他のナンバリング方式も自明であり同等に置換することができる。

#### 【0028】

交換と指定されたSBCONチャネル間の結合はブリッジユニット107が使用して維持することもできる。各情報単位201は各フレーム202の各FC-2内

の交換ID値により示される特定の交換に属する。ブリッジ装置107は受信したFC-2フレームをSBCONフレームに変換し、交換ID値を使用してSBCONフレームを適切なSBCONコントロールユニット103へ転送する。

#### 【0029】

スイッチ108はディレクタ105内の各SBCONポート106との個別のポイントツーポイントリンクを作り出して管理する。図1には2, 3のリンクしか図示されていないが、典型的なディレクタ105は任意のポート106と任意他のポート106間のリンクを管理することができる(248以上のポート106まで)。FCフレームが受信されて検証されると、フレームを調べてそれが予め確立されている交換(exchange)に関連しているか、あるいは、新しい交換が作り出されているかどうかを確認される。新しい交換に対しては、ポートコントローラ107がスイッチ108へのSBCON接続要求を発生する。SBCON接続要求は、マトリクススイッチユニット108がSBCONプロトコルに従ってチャネルおよび特定のSBCONコントロールユニット103間に接続を確立できるようにする、FC-4ヘッダーから抽出した情報を含んでいる。スイッチ108は接続が確立されたかを示す、あるいは、接続を防止した事情(例えば、話し中もしくは拒否)があればそれを示す応答符号を戻す。一度接続が確立されると、ポートコントロール107は受信したフレームをSBCONフレームに変換し、それは交換の持続時間中指定されたSBCONインターフェイスへ転送される。

#### 【0030】

図3Aは本発明の特定のインプリメンテーションに従ったポートコントロール107をブロック図形式で示す。フロントエンドユニット301が本質的に従来標準に従った方法でFC-0物理的およびFC-1トランスポート機能を実現する。理解を助けるために、本発明の記述はメインフレーム102(図1に示す)に関しており、“TX”で示すフレームおよびパスウェイ(pathways)はメインフレーム102からの送信を表わし“RX”はメインフレーム102へ向けられるデータに関するフレームおよびパスウェイに適用される。フロントエンドユニット301はローカルプロセッサ302もしくはリンクコントローラ303(図3にLC\_0からLC\_7で示す)の1つに接続されるTXフレームハンドラー309を使用

してFC-2フレームを処理する。また、フロントエンドユニット301はローカルプロセッサ302もしくはリンクコントローラ303の1つからRXフレーム発生器311を介して発信されるFC-2フレームを受信する。

#### 【0031】

ローカルプロセッサ302はFCリンク上のFC-標準プリミティブシーケンス (primitive sequences) を制御する機構を含んでいる。プリミティブシーケンス処理は実質的にFC-標準により指定され、本発明を完全に理解するのに詳細に理解する必要がない。これらの技術をさらに理解するには出版されているファイバチャネル標準を参照するのが適切である。本発明の目的のためには、フロントエンドユニット301がローカルプロセッサ302と協働して、FCフレームを含む任意のデータが通される前に、FCリンクを“アクティブ”状態までシーケンスするように動作することを言及すれば十分である。

#### 【0032】

TXフレーム（すなわち、メインフレーム102から受信されたフレーム）に関して、有効なFC-4データ以外の任意のFCフレームもしくはFC-4リンクフレームがローカルプロセッサ302へ転送される。それはFC-2フレームヘッダー内のルーティング制御 (R\_CTL) ビットにより決定される全てのFC-2フレームおよびクラス3デリミッタ（すなわち、“SOFi3”もしくは“SOFn3”にセットされたデリミッタタイプ）以外のSOFフィールドを有する任意のフレームを含んでいる。ポートコントロール107は、フレームがローカルプロセッサ302に向けられるかあるいはリンクコントローラ303の1つに向けられるかに応じて、受信したフレームを異なるように処理する。特に、ローカルプロセッサ302に向けられたいかなるフレームもフレーム単位 (frame-by-frame basis) で処理され、リンクコントローラ303に向けられたフレームはシーケンス境界（すなわち、情報単位）ベース (sequence boundary basis) で処理される。

#### 【0033】

ポートコントロール107はさまざまなデータを格納するためのいくつかのレジスタ304を含んでいる。いくつかのレジスタ304は、その中に含まれるデ

ータがポートコントローラ自体もしくは全てのリンクコントローラ303に関連する点でグローバルであり、他のレジスタはリンクコントローラ303の特定の1つに対応する。好ましいインプリメンテーションでは、RXパスは一意的範囲の交換ID (OX\_ID) 番号およびシーケンスID (SEQ\_ID) 値を各リンクコントローラ303へ割り当てることにより実現され、その範囲はレジスタ304に格納される。また、レジスタ304はポートコントローラのアドレスおよびグループベースアドレス (group base address) を格納するのにも使用され、その重要性については後述する。RXデータ (すなわち、ファブリック101へ行くデータ) に対するFC-2フレーム発生プロセスにおいて使用されるレジスタ304内にいくつかの値が格納される。特定アプリケーションのニーズに応えるように、任意の数およびサイズのレジスタ304を設けることができる。

#### 【0034】

操作において、フロントエンド301は受け手識別 (D\_ID) がポートコントローラの格納されたアドレスフィールドに一致するかどうかまたフレームの送り手識別 (S\_ID) がポートコントローラの格納されたグループベースアドレス値に一致するかどうかを調べることによりTXフレーム上のアドレス検証を実施する。このアドレス検証ステップに合格しない場合には、フレームはローカルプロセッサへ転送される。アドレス検証後、ポートコントロール107はFC-2ヘッダーフィールド内の情報に基づいて他のフレーム検証検査を実施することができる。

#### 【0035】

ポートコントローラ107はルックアップテーブル (LUT) 306も含んでいる。LUT 306 (図6に詳細に示す) は各オープン交換に対するスロットを含み、各スロットが複数のフィールドを保持する。好ましいインプリメンテーションでは、LUT 306は特定のオープン交換に対応するOX\_IDフィールドによりTXフレームに対してインデクスされる。好ましいインプリメンテーションでは、LUT 306はリンクコントローラ番号 (すなわち、LC\_0からLC\_7) によりRXフレームに対してインデクスされる。このようにして、特定交換データを格納し、維持し、LUT 306から検索することができる。

## 【0036】

図3Aに示すように、コントローラ107はTXバッファメモリ308およびRXバッファメモリ312を含んでいる。各バッファ308および312が複数の場所（すなわち、ライン、エントリ、スロット）を含んでいる。各場所（location）はバッファ記述子（descriptor）とも呼ばれる。各バッファ記述子は情報単位を保持するサイズとされる。特定の例では、RXバッファメモリ312は30のバッファ記述子を含んでいる。特定の例では、TXバッファメモリ308も30のバッファ記述子を含み、その一部は“カットスルー”（cut-through）記述子として割り当てられ、そのもう1つの部分は“始動”（start up）記述子307として割り当てられる。

## 【0037】

ポートコントローラ107の重要な機能はフレームを特定のリンクコントローラ303に結合してバウンドリンクコントローラ（bound link controller）303へフレームを向けることである。交換結合が確立される前に、フロントエンドユニット301により提供される新しいフレームは交換結合が決定される間格納すなわちバッファされなければならない。始動バッファ307がこの暫時格納を行う。好ましくは、本発明は交換結合が確立されるまでメインフレーム102は同じ交換IDを有する多数のフレームを送ることができないように実現される。始動バッファ記述子307は各オープン交換に対して少なくとも1つの完全なフレーム、すなわち図3Aの特定の実施例では8フレーム、を格納するのに十分な大きさでなければならない。多数の交換結合が同時にセットアップされそうにないある応用では、より小さい始動バッファが可能である。

## 【0038】

各FCフレームが、シーケンスの最初のフレームであるか、シーケンス内の中間フレームであるか、あるいはシーケンス内の最後のフレームであることを示す情報を、FC-2ヘッダー内のそのフローコントロール（F\_CTL）フィールド内に含んでいる。ここで使用されるシーケンスは図2に示す情報単位201に対応する。各TXフレームに対して、TXハンドラー309はFC-2ヘッダーからの交換IDフィールド（OX\_ID）をコピーし、LUT306に質問して交換結合が既

に確立されているかどうかを確認し、このフレームに対して“オープンする”（後述するいかなるエラー条件もない）。このような表示を受信すると、TXハンドラー309は新しいフレームのOX\_IDを使用してLUX306をインデクスし、そのOX\_IDに対する交換が現在オープンすなわちアクティブであるかどうかを確認するために調べる。

#### 【0039】

最初に、交換結合は存在しない。新しい交換の最初のフレームは始動記述子エリア307に格納される。フレームハンドラー309は、それが交換内の最初のフレームであることを示すFCフレームの受信に応答してLUT306内に新しい交換結合を作り出す操作を実施する。それが交換内の最初のフレームであることを示すフレームに遭遇すると、フロントエンド301は特定交換制御データを含むFC-4ヘッダーを探す。

#### 【0040】

図3BはTXバッファメモリ308および始動記述子エリア307を実現するのに使用されるメモリ構造の特定のインプリメンテーションを示す。図3Bのインプリメンテーションにおいて、各々が特定のリンクコントローラ303に関連づけられるように複数のメモリチップ315が構成される。チップセレクト信号が特定のメモリを選択的に活性化させる。メモリチップ315は共通アドレスバスおよびデータバスを共用する。アドレスバス上に供給されたアドレス信号は各チップ315内の特定の場所を示す。

#### 【0041】

好ましいインプリメンテーションでは、FC-4ヘッダーに関連する縦冗長検査(LRC)符号が調べられる。LCRチェックにパスすると、SBCONリンクアドレスがFC-4ヘッダーから抽出されTXハンドラー309は抽出されたSBCONリンクアドレス、格納されたグループポートアドレス、およびグループポート番号を使用してスイッチ108（図1に示す）にSBCON接続要求を行う。スイッチ108は要求を処理し、接続が確認されているもしくは否定されていることを示す応答符号を戻す。スイッチ応答の一部は新しい交換に使用される特定のリンクコントローラ303を、好ましくは、スイッチ108により戻された送り手番号の

下位3ビット内で識別する。次に、ポートコントロール107は受信したフレームのOX\_IDに対応するOX\_IDに対するLUT306内のエントリを更新して、交換がアクティブであることを示し戻されたリンクコントローラIDをLUT306のそのOX\_IDスロット内に格納してLC識別をこのOX\_IDに“結合する”。

#### 【0042】

始動フレームが受信されると、全てのチップセレクトラインがアクティブとなり始動フレームは各チップ315の同じ始動記述子エリア307内に書き込まれるようにされる。この特定の交換に対してリンクコントローラ303への結合が決定されると、特定の始動記述子の表示がそのリンクコントローラに与えられる。このようにして、初期フレームは交換機に関連する適切なリンクコントローラ303へ転送される。次に、交換の初期フレームは始動記述子307からバウンドリックコントローラ303へ転送される。結合プロセスの結果は特定のリンクコントローラ303が交換に関連づけられ、特定のTXバッファメモリ記述子308（“始動”記述子とも呼ばれる）が特定のリンクコントローラ303に関連づけられることである。シーケンス内の中間および最終フレームに対しては、FC-2ヘッダー情報は交換内の最初のフレームのFC-2ヘッダー情報と一致するため必要がない。中間および最終フレームは、初期フレームに対して使用された始動記述子307ではなく、それらの関連するカットスルー記述子308内へ向けられる。カットスルー記述子は環状リスト（circular list）として動作する。既存の交換結合が中間フレームに使用され、F\_CTLフィールドが交換の終りを示す時に既存の交換結合は閉じられるすなわち除去される。

#### 【0043】

前記したように、好ましいインプリメンテーションでは、CRC検査は完全なFC-2フレーム上でフレームが処理される前に実施される。通常ある種の伝送エラーを示す任意のCRCエラーは1つのレジスタ304を介して報告しなければならない。初期フレームに対して、CRC検査は好ましくは結合要求処理が開始される前に実施され、新しい交換結合を作り出すプロセスが開始される前にCRCエラーが検出されるようにされる。CRCエラーが検出される前に交換結合が作り出されると、割込み符号によりローカルプロセッサ302に知らせてロー

カルプロセッサ302がエラー処理手順を実行するようにしなければならない。  
望ましくは、マトリクスコントローラ接続要求に使用したD\_ID値はCRCが検証されるまで保持される。CRCエラーを有する任意のフレームが望ましくは廃棄される。

#### 【0044】

各リンクコントローラ303はメモリを含み、その中にいくつかのTXバッファ記述子およびRXバッファ記述子の実現される。TXバッファ記述子はSBCONフレームを形成するのに必要な制御およびペイロードデータを保持する。同様に、RXバッファ記述子はFRフレームを形成するのに必要な制御およびペイロードデータを保持する。好ましいインプリメンテーションでは、図4に示す、単一のRX/TXバッファ記述子フォーマットが使用される。図4に示す記述子フォーマットは理解を容易にするための特定の語構成を示すが、本発明の教示を逸脱することなくRX/TXバッファ記述子は図4に示すサイズ、構成、および相対的比率を著しく変えられることをお判り願いたい。

#### 【0045】

RX/TXバッファ記述子308および312内に保持される制御情報は記述子制御(DESC\_CTL)、ルーティング制御(R\_CTL)、およびIUが交換内の初期、中間もしくは最終情報単位であるかどうかを示す情報単位(IU\_CTL)を含んでいる。記述子制御フィールドは、例えば、記述子の所有権、状態情報、その他の記述子に固有の管理値(descriptor-specific management value)を示す値を保持するのに使用される。ルーティング制御フィールドおよびタイプフィールド(Type fields)は対応するFCフレームのFC-2ヘッダーからコピーされる。IU\_CTLフィールドは従来のFC-2ヘッダーのフレーム制御(F\_CTL)フィールドに類似しているが、FCフレームではなく“情報単位”に関連する値しか含んでいない。バイトカウントフィールドはペイロードバッファ部内のバイト数を示す値を保持し、したがって、変動するサイズのフレームによってデータの効率的なトランスポート(transport)ができるようにすることができる。

#### 【0046】

FC-4ヘッダーフィールドはその全体をRX/TXバッファ記述子へコピー

一され、残りの領域はFC-4ペイロードデータに対して割り当てられる。標準Fcフレームはおよそ2Kバイトのペイロード最大サイズを有し、標準SBCONフレームはおよそ1Kバイトの最大ペイロードサイズを有することをお判り願いたい。したがって、各バッファ記述子は多数のSBCONフレームもしくはFCフレームに対して十分なペイロードデータを保持することができる。ポートコントローラ107はシーケンスを適切な数のフレームへセグメント化する責務を負う。好ましくは、シーケンスフレームの終りを除く各フレームは最大長FCフレームである。

#### 【0047】

フレーム処理中に、TXバッファ記述子およびRXバッファ記述子内のさまざまなフィールドがペイロードおよびヘッダーデータで充填される。TXフレームハンドラー309はTXバッファメモリ308内にフレームを置くように機能する。リンクコントローラ303はTXバッファ記述子308からのデータを使用してスイッチインターフェイス313へ送信するSBCONフレームを発生する。RXフレーム発生器311はRXバッファ記述子312からのデータを使用してフロントエンド301へ送信するFCフレームを発生する。

#### 【0048】

典型的な交換機は多数の情報単位を含むことが考えられる。同様に、多くの情報単位すなわちシーケンスが多数のFCフレームを含む。各FCフレームは後に1つ以上のSBCONフレームの情報フィールド内へ転送されるある量のペイロードデータを含むが、FCフレームとSBCONフレーム間には1:1の対応がない場合が多い。また、マルチフレームシーケンスに対しては、シーケンスの後続フレームのペイロード部だけが記述の情報フィールドに書き込まれる。シーケンスフレームのFC-2ヘッダー内の重複情報はコピーされない。したがって、TXバッファ記述子に実際にコピーされるデータ量は対応するFC-2フレーム内の実際のデータ量とは異なる。

#### 【0049】

マルチフレームシーケンスに対して、ポートコントローラ107は全体シーケンスにわたって“実際の”FC-4ペイロードのバイトカウントを維持する。“

実際のFC-4ペイロード”はペイロードのSBCON情報フィールド内へ実際にコピーされる部分のことであり、FC-4フレーム内のヘッダー情報はカウントしない。この実際のペイロード値はリンクコントローラが応答できるようにSBCONフレーム内に供給される。ポートコントローラ107はマルチフレームシーケンスを検出し、後続フレームに対して各リンクコントローラ303のメモリ構造内にライトポインターを維持する責任がある。したがって、FC-2ヘッダーのタイプフィールド、R\_CTLフィールドが記述子情報の一部として含まれる。完全なシーケンスがTXバッファ記述子にコピーされていると、ポートコントロールユニット107は完全な記述子の所有権を関連するリンクコントローラに変え、現在の交換結合に対する次のフレームをアSEMBルして処理するための次のTXバッファ記述子にそのポインターを調節する。

#### 【0050】

ポートコントロールユニット107は新しいシーケンスを受信したら、それを使用する前にそれがカットスルーバッファを所有することを検証しなければならない。ポートコントロールユニット107がバッファ記述子を所有せず（例えば、所有権が既にリンクコントローラ303に転送されている）かつ新しいシーケンスが受信される場合には、オーバラン（overrun）状態が存在しレジスタ304のエラー状態レジスタ内に表示しなければならない。ポートコントロールユニット107はその交換上の全ての後続シーケンスに応答して消失（missing）シーケンスエラー符号を発生しなければならない。

#### 【0051】

RXフレーム（例えば、ファブリック101へ行く）処理に対して、ポートコントローラ107は各リンクコントローラ303もしくはローカルプロセッサ302により作り出されたRXバッファ記述子からFC-2フレームを作り出して送信する責任がある。ポートコントローラ107はFC-2ヘッダーの生成を助ける各リンクコントローラ303に対する1組のレジスタだけでなく、レジスタ304内のいくつかのグローバルレジスタを含んでいる。グローバルレジスタは、前記したように、FC-2ヘッダーのS\_IDおよびD\_IDのために使用されるポートコントローラのアドレスおよびグループベースアドレスを格納するレジスタ

を含んでいる。グローバルレジスタはエラー状態、および最大フレームサイズ等のグローバルパラメータを保持するためのレジスタも含んでいる。

#### 【0052】

R Xフレーム処理に対する重要な機能はOX\_IDの割当ておよびSBCONコントロールユニット103より発信されるフレームに対する連番付けを管理することである。このリンクコントローラ303に対する開始OX\_IDを表わすOX\_ID値を保持するのに特定リンクコントローラレジスタが使用される。このリンクコントローラ303に対するOX\_ID数を表わす値を保持するのにもう1つの特定リンクコントローラレジスタが使用される。他のレジスタは開始シーケンスID値および関連するリンクコントローラに対する最大シーケンス数を示す値を保持する。特定応用のニーズに応えるために他のレジスタを使用することができる。

#### 【0053】

FC-2ヘッダーの残りはR\_CTL, およびR Xバッファ記述子の情報単位制御(IU\_CTL) フィールドから作り出される。IU\_CTLはF\_CTLの上位ビットを発生するために使用され、IUが交換内の最初のIUであるか交換内の最後のIUであるかを示す。交換結合が作り出されると、交換の持続時間中シーケンスを識別するFC-2SOFフィールドに対するF\_CTLビットおよびデリミッタタイプを作り出すのはポートコントローラ107の責任である。各リンクコントローラに対するSEQ\_CNT値がレジスタ304の1つに格納され、好ましいインプリメンテーションにおいて送られる各フレームにより増分するフリーランニングカウンタとして実現される。SEQ\_CNT値は各シーケンスがSEQ\_CNT=0で開始するようにシーケンスの終りにゼロへリセットされる。このようにして、ゼロに等しくないがさもなくばシーケンス最初のフレームとして現れるSEQ\_CNTを示すフレームに到達する時にエラー状態（例えば、消失フレーム）を検出することができる。

#### 【0054】

LC303はFC-4シーケンスを完了すると、関連するR X記述子の所有権をポートコントローラ107へ渡してIUが準備完了であることを示す。R Xバッファメモリ312内のR X記述子はラウンドロビン式にポーリングされて、送る準備が完了しているIUを識別する。準備完了であるIUが識別される度に

、バッファ記述子所有権ビットが検証されバッファ記述子バイトカウント値が読み出される。FC-2ヘッダー情報が読み出され、フレーム発生器311によりFC-2フレームが作り出されFCリンクを介して送信される。フレーム発生器311がシーケンスを完了すると、ポートコントローラ107はその記述子の所有権をリンクコントローラ303へ変える。

#### 【0055】

図5は本発明に従った機構および方法の操作を例示するデータフロー図を示す。ソフトウェアアプリケーション501が図1に示すメインフレーム102を実行する。ソフトウェアアプリケーション501はメインフレーム102に関連するもしくは一体化されたN\_Port 502へ送られる情報単位（“IU”で示す）を発生する。ソフトウェアアプリケーション501は発信交換識別（OX\_ID）を割り当てる。このOX\_IDは、所与の交換を識別する一意的OX\_ID値が選択され、その交換に関連する全ての情報単位に割り当てられたOX\_IDが付されるように、従来の方法で割り当てられる。OX\_IDを割り当てるのに、単純な逐次割当てを含む、任意の機構を使用することができる。

#### 【0056】

IUは従来の方法でFCフレーム内で送られ、ポートコントローラフロントエンド301へ転送される。フロントエンド301はFC-2ヘッダー情報を剥ぎ取って（strips off）情報単位を再度作り出す。次に、情報単位は適切なリンクコントローラへ転送され、それは前記した交換結合方法論に従って決定される。新しいフレームに対して、受信リンクコントローラ303はメインフレーム102への応答フレームを作り出すのに必要なFC-2ヘッダーからの情報を含む“制御ブロック”を発生する。制御ブロックは特定の例においてLUT306内にフィールドとして格納される。1つ以上のSBCONフレームが発生されSBCONコントロールユニット103へ送られる。

#### 【0057】

典型的に、SBCONコントロールユニットはポートコントロールユニット107へ返送されるACCEPTのような応答フレームを発生する。受信したSBCONパケットのS\_IDがグループベースアドレスにセットされる。SBCONコントロールユニット

と特定のリンクコントローラ303間の回線接続は交換の持続時間中確立されたままであるため、SBCON応答は正しいリンクコントローラ303により受信される。

#### 【0058】

リンクコントローラは前に保存された制御ブロック内に格納されたフレーミングデータ (framing data) を使用してSBCONフレームを情報単位へ再フォーマット化する。リンクコントローラ303は、フレームがオープン交換に関連しているかあるいは新しい交換を作り出す必要があるかに関する、フロントエンドユニット301への表示を含んでいる。新しい交換に対して、リンクコントローラはその割り当てられた範囲からのOX\_ID値を割り当てる。既存の交換に対して、情報単位にはこの交換を識別する一意的OX\_ID値が付される。フロントエンドユニット301はファブリック101を介してメインフレーム102返送されるFC-2フレームを作り出す。

#### 【0059】

ソフトウェアアプリケーション501は、受信したフレームのFC-2ヘッダー内に格納されたOX\_ID値に基づいて、このフレームが同じOX\_IDを有する前に発行された情報単位に関係するかどうかを決定する。この点において、交換対 (exchange pair) はオープンでありパケットはファブリック101の両端間で送信し合うことができ、SBCONコントロールユニット103により整然と受信することができる。発信および着信交換が共にエラーもしくは時間切れ状態により明確にもしくは暗黙に終端されるまで、交換対はオープンのままである。

#### 【0060】

重要なのは、ソフトウェアアプリケーション501が多数の同時交換を確立し、したがって、最初のSBCON接続に対する応答を待機しながら他のSBCON接続へ他のパケットを送ることができることである。したがって、8つまで恐らくはそれ以上の交換が同時にオープンとなって、各交換が独立した200 Mbits/秒チャネルヘデータを供給することができる。ノードポートの相関関係から、単一の交換がそれ自体とフロントエンドユニット301との間でオープンとされ、全てのパケットがその交換へアドレスされる。

## 【0061】

図6はLUT306に対する単純化された構造を示す。LUT306は論理テーブルデータ構造もしくはOX\_IDエントリによりインデクスされるコンテンツアドレス可能メモリ構造として実現することができる。各スロット601が複数のエントリすなわちフィールドを含んでいる。状態情報は関連する交換がクローズド、オープン（すなわち、マトリクススイッチユニット108への要求が保留されている）およびアクティブ（すなわち、結合が確立されている）であるかどうかを示す。“結合”フィールドはこの交換に結合される特定のリンクコントローラ303を示すグループ番号識別を保持する。SEQ\_IDフィールドはこの交換上で受信した最後のフレームからのFC-2シーケンスIDを保持する。SEQ\_FLGフィールドはこの交換上のシーケンスがアクティブであるか終端しているかを示す。マルチフレームシーケンスに対して、SEQ\_CNTフィールドはそれが受信した最後のFC-2シーケンスであるかどうかを示し、R\_CTLフィールドはシーケンスの最初のフレームからのFC-2 R\_CTLの情報カテゴリフィールドを保持する。エラーフィールドは現在シーケンス上でエラーが検出されているかどうかを示し、またエラーのタイプを示す符号化された値を保持することができる。

## 【0062】

特定のインプリメンテーションでは、初期化すなわちブートアップ中に、ローカルプロセッサ302がポートコントローラ107内のリンクコントローラの状態検査を行って、それらの動作状態を決定する。特定のポートコントローラ107には8つよりも少ないLC303がインストールされていることがあり、あるいは、さまざまな理由から1つ以上のLC303が動作しないことがある。特定のインプリメンテーションでは、1つのオープン交換が各動作リンクコントローラ303、すなわち、図3Aの例では最大8つ、に割り当てられる。ローカルプロセッサ302は許されるオープン交換対数を示すFC-4ブリッジ制御フレームを発生する。

## 【0063】

メインフレーム102（図1に示す）はこれらの交換クレジット（exchange credits）をポートコントローラ107から受信するまで、いかなるパケットも交

換へ送ることができない。言い換えれば、ポートコントローラ107はクローズしている交換に対応するOX\_IDを有して送られたいかなるパケットも棄却する。新しい交換を確立する必要があるパケットを送る時は、メインフレーム102は交換クレジットを断念する。メインフレーム102がその交換クレジットを全て断念してしまうと、既存の交換がクローズするまでどの新しい交換にもメッセージを送ることができない。交換がクローズされると、ポートコントローラ107はFC-4ブリッジ制御メッセージを発生して現在クローズしている交換に対応する交換クレジットをメインフレーム102に与える。

#### 【0064】

交換クレジットは初期化時だけでなく実行時にダイナミックに決定できることが考えられる。ローカルプロセッサ302は、もう1つのFC-4ブリッジ制御フレームをメインフレーム102に送ってその特定のリンクコントローラに対する交換クレジットを除去することにより、ローカルコントローラ303をダウンさせるすなわち消勢することができる。

#### 【0065】

動作中に、好ましいインプリメンテーションでは、そのシーケンスの初期フレームが受信され交換が確立されているという受信応答(acknowledgement)を受信するまで、チャネルはOX\_IDへフレームを送れないことが必要である。この特徴により始動時にTX記述子バッファがオーバランするのを回避する。チャネルが多数のメッセージを同じ交換番号へ送ることを許される場合には(同じOX\_IDを使用して)、フレームレートはポートコントローラ107がそれらを処理できるよりも高速とすることができる。

#### 【0066】

全ての新しい出交換(outbound exchange)上に、即ち、メインフレームとその交換に結合されるリンクコントローラ303間の全ての新しい変換上に同期点を作らなければならない。それは出交換の状況、すなわち、要求されたされないもしくは“船舶通過”シナリオに無関係に(regardless of the context of the outbound exchange, i.e., solicited vs. Unsolicited or any "ship passing" scenarios)要望される。この要求条件は新しい出交換の処理に関連する状態

(latency) によるものである。OX\_IDとリンクコントローラ303間の結合がなされていることをメインフレーム102に表示し戻す他の機構はなにもない。ポートコントローラ107によりそのOX\_IDに対して結合が確立されるまで、そのOX\_IDを介したいかなる後続シーケンスの配送も保証されない。

#### 【0067】

本発明をある程度詳細に記述し例示してきたが、本開示は単なる例としてなされるものであり、当業者ならば後述する本発明の精神および範囲を逸脱することなく部品の組合せおよび配列をさまざまに変更できることをお判り願いたい。

#### 【図面の簡単な説明】

##### 【図1】

本発明に従った機構および方法が実現される分散形計算環境を示す図である。

##### 【図2】

本発明に従った情報単位の記述を示す図である。

##### 【図3A】

本発明の特定のインプリメンテーションに従ったポートコントローラを示すブロック図である。

##### 【図3B】

本発明の特定のインプリメンテーションに従った図3Aに示すポートコントローラの一部を示すブロック図である。

##### 【図4】

本発明に従ったインプリメンテーションにおいて使用される第1のデータ構造を示す図である。

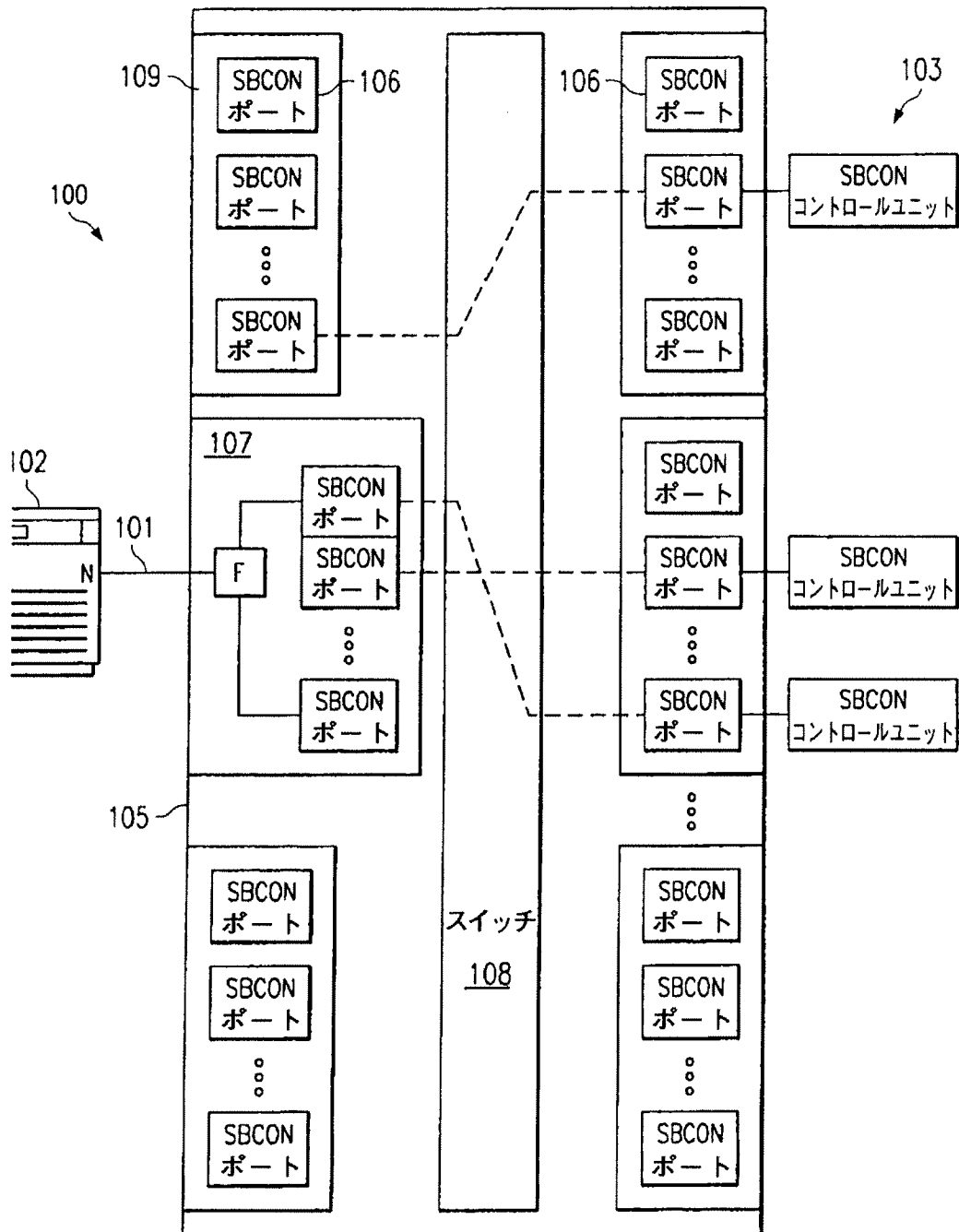
##### 【図5】

本発明に従った機構および方法の操作を示すデータフロー図である。

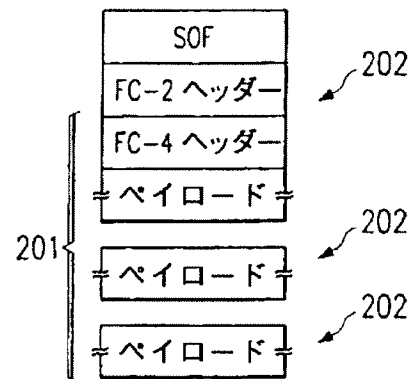
##### 【図6】

本発明に従ったインプリメンテーションにおいて使用される第2のデータ構造を示す図である。

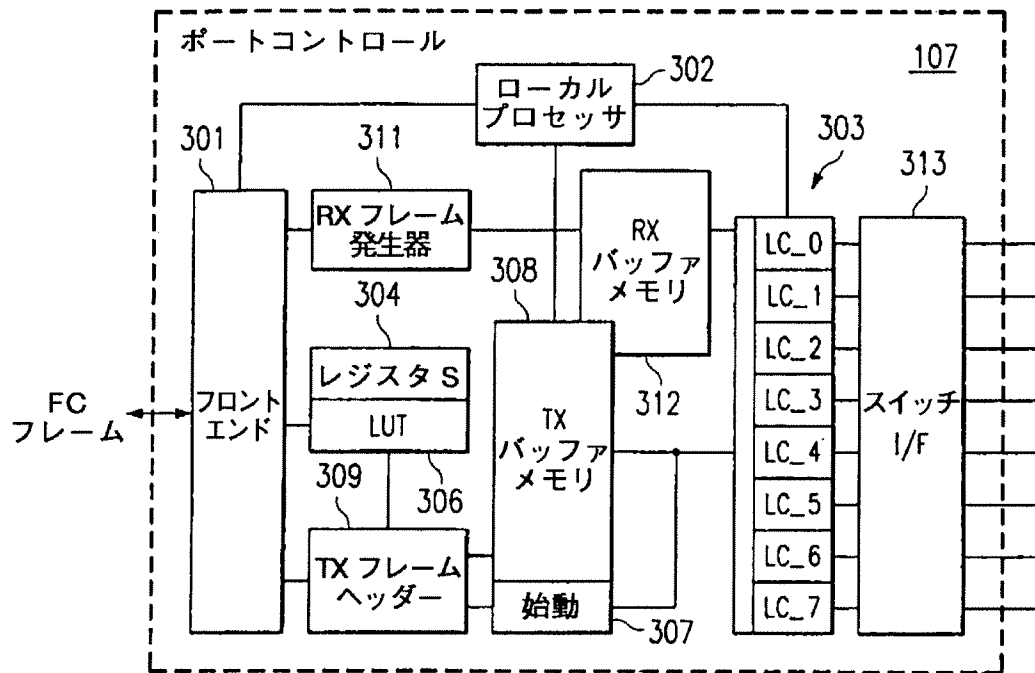
【図1】



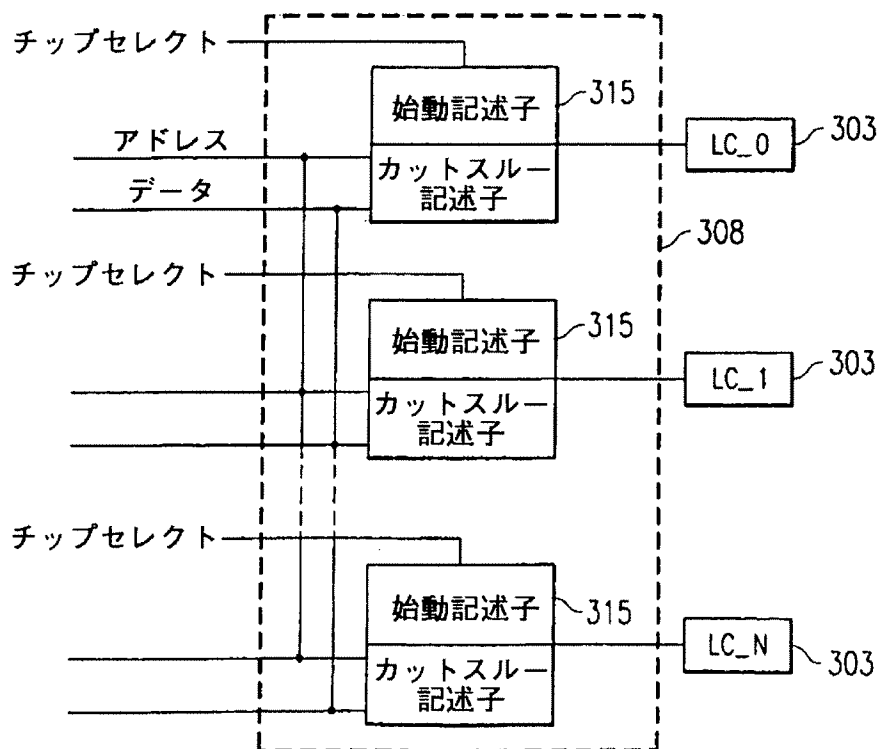
【図2】



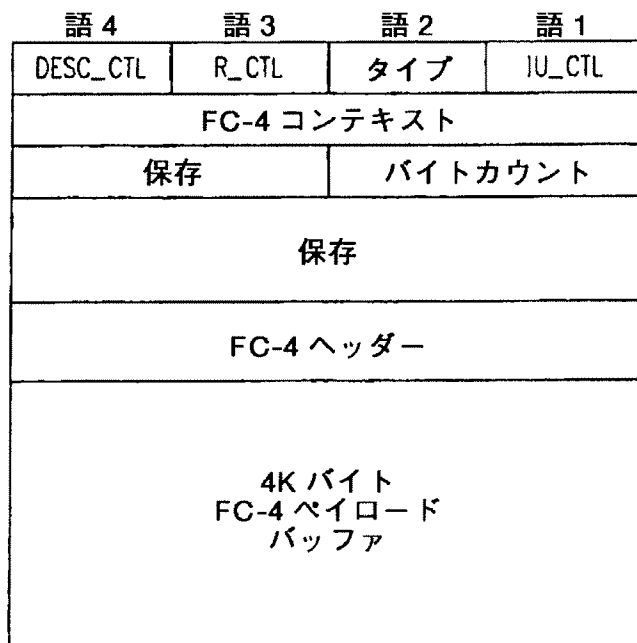
【図3A】



【図3B】



【図4】





【図 6】

601

LUT										306
OX_ID_01	状態	結合	SEQ_ID	SEQ_FLG	SEQ_CNT	R_CTL	エラー			
OX_ID_02	状態	結合	SEQ_ID	SEQ_FLG	SEQ_CNT	R_CTL	エラー			
OX_ID_03	状態	結合	SEQ_ID	SEQ_FLG	SEQ_CNT	R_CTL	エラー			
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮			
OX_ID_N	状態	結合	SEQ_ID	SEQ_FLG	SEQ_CNT	R_CTL	エラー			

【手続補正書】特許協力条約第34条補正の翻訳文提出書

【提出日】平成13年5月2日(2001. 5. 2)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】 データ通信アーキテクチャであって、

第1のデータレートの通信をサポートする入出力(I/O)ポートを有する複数の装置と、

いくつかのI/Oポートを有するデータプロセッサであって、各I/Oポートが第2のデータレートの通信をサポートし、第2のデータレートは第1のデータレートの少なくとも2倍であるデータプロセッサと、

データプロセッサI/Oポートの1つに接続されて第2のデータレートをサポートする通信リンクと、

通信リンクおよび複数の装置のI/Oポートに接続されたブリッジ装置であって、第2のデータレートの通信リンクを第1のデータレートの複数の通信リンクへ変換するブリッジ装置と、

を含むデータ通信アーキテクチャ。

【請求項2】 請求項1記載のデータ通信アーキテクチャであって、第1のデータレートの複数の通信リンクは互いに実質的に独立しているデータ通信アーキテクチャ。

【請求項3】 請求項1記載のデータ通信アーキテクチャであって、複数の装置のI/Oポートは回線交換通信をサポートし、

データプロセッサのI/Oポートはパケット交換通信をサポートする、データ通信アーキテクチャ。

【請求項4】 請求項1記載のデータ通信アーキテクチャであって、複数の装置のI/OポートはSBCONをサポートするデータ通信アーキテクチャ。

【請求項5】 請求項1記載のデータ通信アーキテクチャであって、データプロセッサのI/Oポートはファイバチャネルをサポートする、データ通信アーキテクチャ。

【請求項6】 請求項1記載のデータ通信アーキテクチャであって、データプロセッサは複数のメッセージを発生し、各メッセージが複数の装置の特定の1つを有するバーチャルチャネルに対応するデータ通信アーキテクチャ。

【請求項7】 請求項6記載のデータ通信アーキテクチャであって、さらに、複数のバーチャルチャネルを多重化して単一通信リンクとするように接続されたデータプロセッサ内のマルチプレクサを含むデータ通信アーキテクチャ。

【請求項8】 請求項1記載のデータ通信アーキテクチャであって、ブリッジは、さらに、

フレームプロセッサとして動作するフロントエンドと、  
通信リンクを初期化するように動作するローカルデータプロセッサと、  
複数の装置の各々に対するリンクコントローラを含むリンクコントローラユニットであって、各リンクコントローラが複数の装置の対応する1つとの通信チャネルをサポートするリンクコントローラユニットと、  
を含むデータ通信アーキテクチャ。

【請求項9】 請求項8記載のデータ通信アーキテクチャであって、さらに、

データプロセッサへ交換クレジットメッセージを発生するように動作するローカルプロセッサ内の第1の交換クレジット機構と、  
交換クレジットメッセージを受信するように動作するデータプロセッサ内の第2の交換クレジット機構と、  
を含むデータ通信アーキテクチャ。

【請求項10】 請求項8記載のデータ通信アーキテクチャであって、第1の交換クレジット機構はリンクコントローラユニット内のいくつかの動作リンクコントローラに応答するデータ通信アーキテクチャ。

【請求項11】 請求項8記載のデータ通信アーキテクチャであって、第1の交換クレジット機構は動作リンクコントローラに関連する複数の装置の組合

せデータレートに応答するデータ通信アーキテクチャ。

【請求項12】 データ接続の管理方法であって、

第1のメッセージを発生するステップと、

第1のメッセージを符号化して第1のデータグラム(datagram)とするステップであって、第1のデータグラムはパケット交換リンクに使用されるメタデータ(meta-data)を含み、第1のメッセージはそれが属する論理交換を一意的に識別する値を保持する交換識別フィールドを含みまた第1のメッセージはそれが属する論理交換を一意的に識別する値を保持する交換識別フィールドを含むステップと、

パケット交換リンクを介して第1のデータグラムを運ぶステップと、

リンクに接続された中間データトランスポート機構内のパケット交換リンクから第1のデータグラムを受信するステップと、

中間データトランスポート機構内にメタデータを格納するステップと、

第1のメッセージを再符号化して第2のデータグラムとするステップであって、第2のデータグラムは回線交換リンクに使用されるメタデータを含むステップと、

回線交換リンクを介して第2のデータグラムを運ぶステップと、  
を含む方法。

【請求項13】 請求項12記載の方法であって、さらに、

回線交換リンクから第2のデータグラムを受信し、

第2のデータグラムに応答して第2のメッセージを発生し、

第2のメッセージを符号化して回線交換リンクに使用されるメタデータを含む第3のデータグラムとし、

回線交換リンクを介して中間データトランスポート機構へ第3のデータグラムを運び、

第2のメッセージを再符号化して第4のデータグラムとし、第4のデータグラムはパケット交換リンクに使用される格納されたメタデータからコピーされたメタデータを含む、

方法。

【請求項14】 請求項12記載の方法であって、さらに、

中間データトランスポート機構内のパケット交換リンクから第1のデータグラムを受信した後で、論理交換の持続時間中交換識別値を特定の回線交換リンクに結合する、ステップを含む方法。

【請求項15】 通信リンクに対するブリッジ回路であって、

全二重パケット交換リンクをサポートするパケット交換側と、

いくつかの全二重回線交換リンクをサポートする回線交換側と、

パケット交換フレームを特定の1つの回線交換リンクに結合する論理結合記述を格納する記憶スペースを有するブリッジ回路内の結合機構であって、ブリッジ回路はパケット交換リンクを介して受信したパケット交換フレーム内に表示された論理交換を識別し、論理交換の持続時間中ずっと論理結合を維持する結合機構と、

受信したパケット交換フレームから選択されたヘッダー情報を保持する結合機構内の記憶構造と、

格納されたヘッダー情報を使用して受信した回線交換フレームをパケット交換フレームへ再フォーマット化する記憶構造に接続されたフレーム発生器と、

を含むブリッジ回路。

【請求項16】 通信リンクの操作方法であって、

高帯域幅接続および複数の低帯域幅接続をサポートするブリッジユニットを提供するステップと、

低帯域幅接続の動作性（operability）を検証するステップと、

動作可能な低帯域幅接続数に基づいて交換機クレジット値を決定するステップと、

高帯域幅接続を介してクレジット値を含むメッセージを発行するステップと、

高帯域幅接続に接続された任意の装置に、その装置から高帯域幅接続を介してブリッジユニットにより通信が受け入れられる前に、少なくとも1つの交換機クレジットを持つことを要求するステップと、

を含む方法。

【請求項17】 請求項16記載の方法であって、検証はブリッジ回路の初

期化中に実施される方法。

【請求項18】 請求項16記載の方法であって、検証は実行時間に動的に実施される方法。

【請求項19】 請求項16記載の方法であって、さらに、

高帯域幅接続に接続された装置からブリッジユニット内のメッセージを受信するステップであって、メッセージは交換クレジットおよび交換識別子を有するステップと、

交換識別子を低帯域幅接続の選択された1つに結合することにより論理交換をオープンとするステップと、

同じ交換識別子を有するブリッジユニットにより受信された後続メッセージをその交換識別子に結合される低帯域幅接続へルーティングするステップと、  
を含む方法。

## 【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US00/08307		
<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC(7) : G06F 13/00 US CL : 710/60, 62; 370/352, 353, 389, 391, 400, 401, 402, 403, 466 According to International Patent Classification (IPC) or to both national classification and IPC				
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) U.S. : U.S. : 710/60, 62; 370/352, 353, 389, 391, 400, 401, 402, 403, 466 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) APS				
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>				
Category *	Citation of documents, with indication, where appropriate, of the relevant passages	Relevant to claim No.		
Y	US 5,712,850 A [ELIA et al] 27 January 1998, col. 1, lines 44-53, col. 3, lines 11-35, 46-58, col. 4, lines 14-25.	1-7, 8-11, 16-18		
Y	US 5,781,547 A [WILSON] 14 July 1998, col. 2, lines 15-31, 43-55, col. 31-14, col. 6, lines 54-67.	8-11, 16-18		
Y, P	US 5,930,250 A [KLOK et al] 27 July 1999, col.2, lines 56-65, col. 7, lines 29-62, col. 8, lines 8-23.	12-15, 19-22		
Y	US 5,819,045 A [RAMAN et al] 06 October 1998, col.3, lines 9-39, col. 4, lines 3-27, col. 8, line 52 to col. 9, line 13.	12-15, 19-22		
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.				
<table border="0"> <tr> <td> <b>Special categories of cited documents:</b>            "A" document defining the general state of the art which is not considered to be of particular relevance            "E" earlier document published on or after the international filing date            "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)            "O" document referring to an oral disclosure, use, exhibition or other means            "P" document published prior to the international filing date but later than the priority date claimed         </td> <td>           "I" later document published after the international filing date or priority date and not in conflict with the application but used to understand the principle or theory underlying the invention            "N" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone            "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is considered with one or more other such documents, such combination being obvious to a person skilled in the art            "X" document member of the same patent family         </td> </tr> </table>			<b>Special categories of cited documents:</b> "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"I" later document published after the international filing date or priority date and not in conflict with the application but used to understand the principle or theory underlying the invention "N" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is considered with one or more other such documents, such combination being obvious to a person skilled in the art "X" document member of the same patent family
<b>Special categories of cited documents:</b> "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"I" later document published after the international filing date or priority date and not in conflict with the application but used to understand the principle or theory underlying the invention "N" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is considered with one or more other such documents, such combination being obvious to a person skilled in the art "X" document member of the same patent family			
Date of the actual completion of the international search 08 AUGUST 2000		Date of mailing of the international search report 29 AUG 2000		
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-2230		Authorized officer THOMAS C. LEE Telephone No. (703) 305-9717		

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/08307

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐

The additional search fees were accompanied by the applicant's protest.

☐

No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/08307

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING  
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-11, drawn to transfer rate regulation.

Group II, claim(s) 12-15, drawn to combined circuit switching and packet switching.

Group III, claim(s) 16-18, drawn to bridge or gateway between networks.

Group IV, claim(s) 19-22, drawn to assignment of variable bandwidth.

---

フロントページの続き

(81)指定国 EP(AT, BE, CH, CY,  
DE, DK, ES, FI, FR, GB, GR, IE, I  
T, LU, MC, NL, PT, SE), JP

(71)出願人 インターナショナル ビジネス マシーン  
ズ コーポレイション  
アメリカ合衆国, 10504 ニュー ヨーク  
州, アーモンク, ニュー オーチャード  
ロード

(72)発明者 ネルソン、ジェフレイ、エル  
アメリカ合衆国 コロラド、ルイヴィル、  
サンランド ストリート 415

(72)発明者 フラオアオス、ゲイリイ、アール  
アメリカ合衆国 コロラド、ブルームフィ  
ールド、 アイリス サークル 1338

Fターム(参考) 5K033 AA04 AA09 BA05 CB08 DA05  
DB19